Preliminary assessment of HRRR reflectivity and echo-top diagnostics

Steve Weygandt, Curtis Alexander, Ming Hu
NOAA ESRL GSD for FAA AWRP MDE

1. INTRODUCTION

This is a report on initial very preliminary assessment of the HRRR reflectivity and echo-top diagnostics associated with the updated Thompson microphysics within the recently frozen GSD real-time experimental RAP / HRRR system. The RAP / HRRR system was frozen in early April and includes numerous enhancements to RAP and HRRR model and data assimilation systems. A complete listing of the system enhancements can be found at:

http://ruc.noaa.gov/pdf/ESRLRAPHRRRchanges2013.pdf

Key systems enhancements include

RAPv2 MODEL
-- updated from WRF-ARW version 3.3.1 base to version 3.4.1 base
-- upgraded PBL parameterization from MYJ to enhanced MYNN
-- Improved Smirnova LSM with 9 levels and improved roughness length specification
-- Other small adjustments to RAP physics modules and radar reflectivity diagnostic code

RAPv2 DATA ASSIMILATION
-- upgraded GSI from standard 3DVAR formulation to EnKF-hybrid assimilation using GFS 80-member ensemble
-- improved specification / removal of precipitation hydrometeors including number concentration for consistency with Thompson microphysics scheme
-- improved soil temperature and moisture adjustment to increase magnitude and make it symmetric for warm/drying and cooling/moistening
-- improved snow cycling -- added snow building capability, removed temperature dependence for snow trimming
-- improved cloud building from METARs for cases with nearby fractional cloud coverage

HRRR MODEL
-- updated from WRF-ARW version 3.3.1 base to version 3.4.1 base
-- upgraded PBL parameterization from MYJ to enhanced MYNN
-- upgraded shortwave radiation scheme from Dudhia to Goddard
-- Improved Smirnova LSM with 9 layers and improved roughness length specification
-- Other small adjustments to RAP physics modules and radar reflectivity diagnostic code

HRRR DATA ASSIMILATION

-- Added 1-h pre-forecast spin-up period including 4 applications of DFI reflectivity-based latent heating

-- Added application of GSI analysis on HRRR domain following 1-h pre-cycle to incorporate latest observations

These changes have added important new capabilities to the RAPv2/HRRR system and have addressed a wide range of shortcomings. Some specific areas of improvement for some of the most key changes include:

As part of our development, we examined a large variety of verification statistics (upper-air, surface, precipitation, reflectivity, and echo-top) for numerous retrospective runs and real-time parallel cycles. Based on this analysis, we implemented those components that showed clear positive impact alone and in conjunction with other system improvements. In addition to this extensive verification effort (crucial for the upgrade decision-making), we will be conducting ongoing real-time verification of the new system against the previous version. In this report, I will present some selected results from the retrospective verification and real-time assessment of the new RAP / HRRR system (no real-time comparison against the previous system yet).

2. Retrospective Verification

Quantitative verification of numerous retrospective experiments was completed as part of the developmental process. We show here reflectivity verification for two retrospective experiments compared against the 2012 real-time forecasts, for a two period (30-31 May 2012) from our 12-day retrospective run. In all plots shown in this section, the red curve represents the real-time run from 2012, the blue curve represents the new system with all changes except the HRRR data assimilation (1-h pre-forecast with 4 15 min. periods or radar reflectivity-based latent heating and the HRRR GSI analysis). The black curve represents the final system including the HRRR data assimilation.

Figure 1 shows the Critical Success Index (CSI) skill scores as a function of forecast length for the three different HRRR cycles described above. As can be seen, even without the HRRR data assimilation, the 2013 changes to the RAP-HRRR (blue curve) yield a significant skill-score improvement through the forecast, compared to the 2012 real-time experimental GSD RAP / HRRR version. Further improvement for the first ~ 4 hr of the forecast period is seen from addition of the HRRR data assimilation system (black curve). An important aspect of this small-scale HRRR data assimilation is that the short-range improvement does not occur at the expense of skill in the longer-range portion of the forecast. One other aspect of the HRRR 3-km assimilation is the dramatic improvement in the 0-h reflectivity analysis.

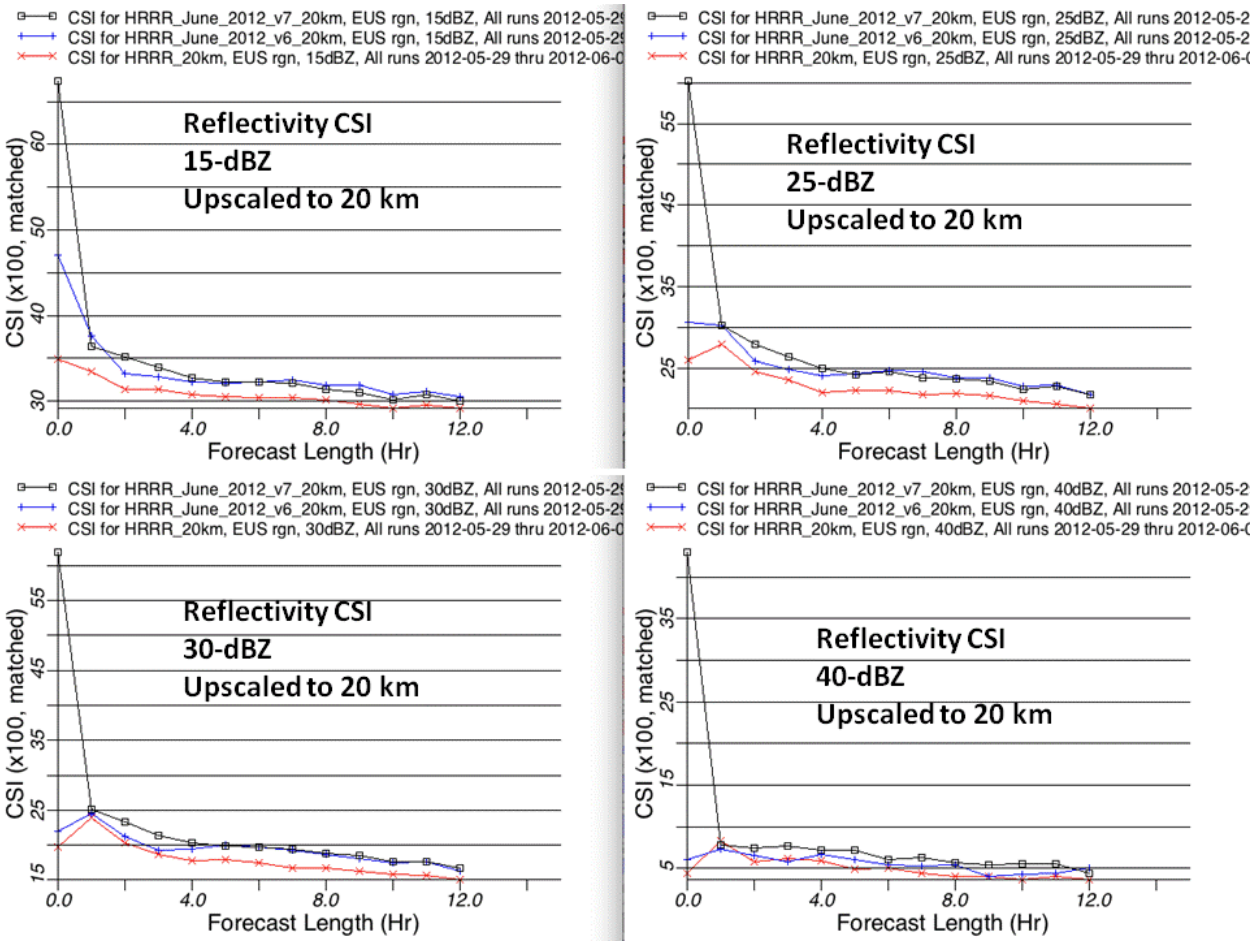


Fig. 1 Critical Success Index (CSI) skill scores as a function of forecast length for three sets of HRRR experiments: red curve – 2012 real-time HRRR, blue curve – retrospective run of 2013 RAP / HRRR with all changes except the HRRR 3-km data assimilation, black curve - retrospective run of the 2013 RAP / HRRR with all changes including the HRRR 3-km assimilation.

Figure 2 shows the HRRR bias as a function of forecast length for the same three HRRR cycles described above. The retro with the new 2013 RAP / HRRR changes excluding the 3-km assimilation (blue curve) shows an interesting pattern with higher bias for mid-range values of reflectivity (25 and 30 dBZ), but a lower bias for low reflectivity thresholds, and especially for high reflectivity thresholds. Addition of the HRRR data assimilation system (black curve) further increases the short-range bias, but has very little impact on the bias for the longer term forecasts.

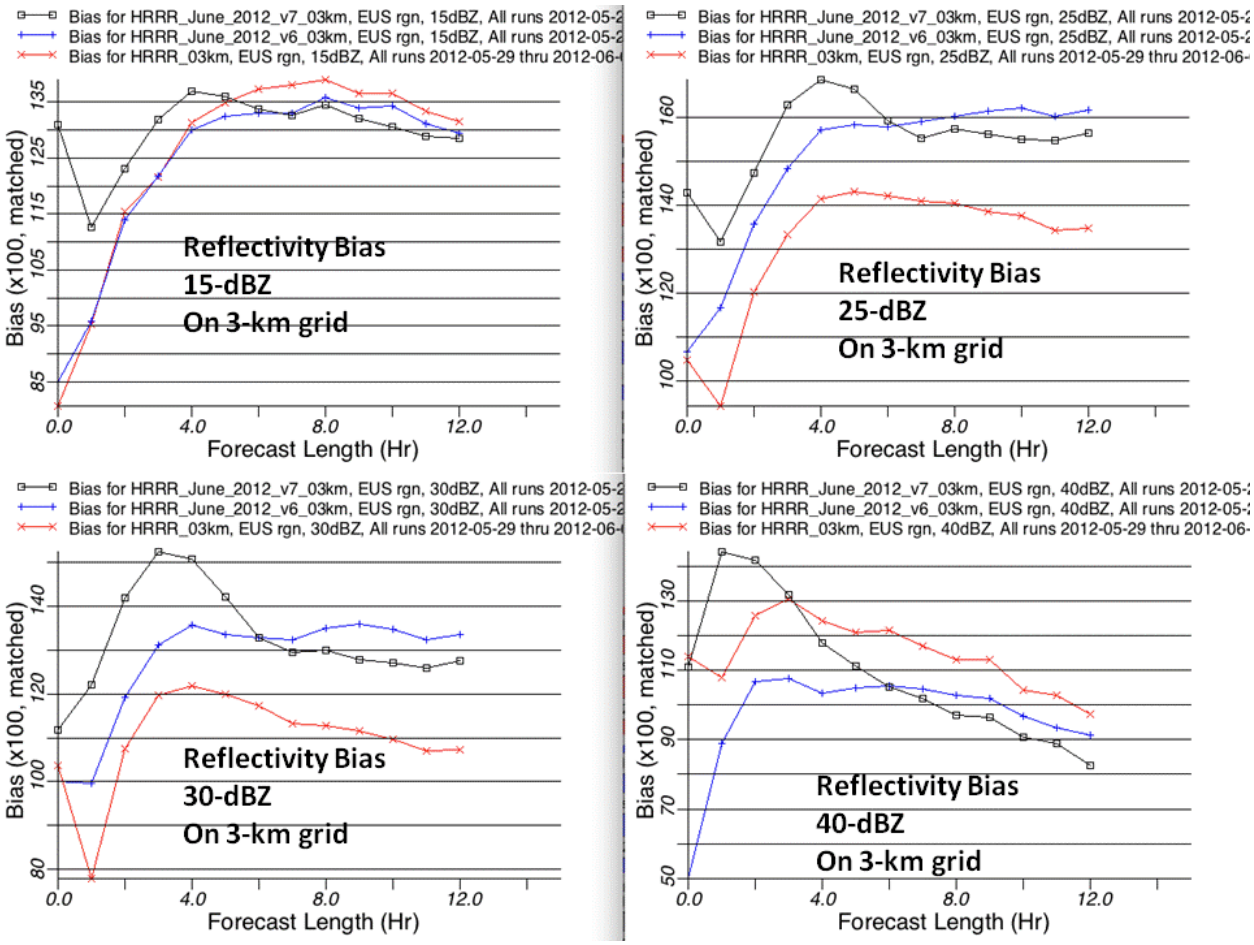


Fig. 2 Bias scores as a function of forecast length for three sets of HRRR experiments: red curve – 2012 real-time HRRR, blue curve – retrospective run of 2013 RAP / HRRR with all changes except the HRRR 3-km data assimilation, black curve - retrospective run of the 2013 RAP / HRRR with all changes including the HRRR 3-km assimilation.

2. Real-time Verification

In addition to the extensive verification of the numerous retrospective runs, we have ongoing real-time verification of the real-time runs. We have a short-term evaluation of the fully configured 2013 GSD real-time HRRR. We are configuring developmental versions of the RAP and HRRR to run with the 2012 real-time set-up, and will supplement the current statistics with a more illuminating comparison in an update to this report next month. Figures 3 and 4 show the CSI and bias for the currently frozen 2013 configuration of the GSD real-time experimental HRRR. While direct comparisons between Figs. 3 and 1, and Figs. 4 and 2 are not possible because of the different case study periods, in general Figs. 3 and 4 mirror the curves in in Figs. 1 and 2, indicating the full real-time system is performing as expected from the retrospective experiments. Note that in Figs. 3 and 4, the full range of abscissa value from 0 through the maximum is shown (this is omitted in Figs 1 and 2, explain why certain aspects to the curves seem accentuated).

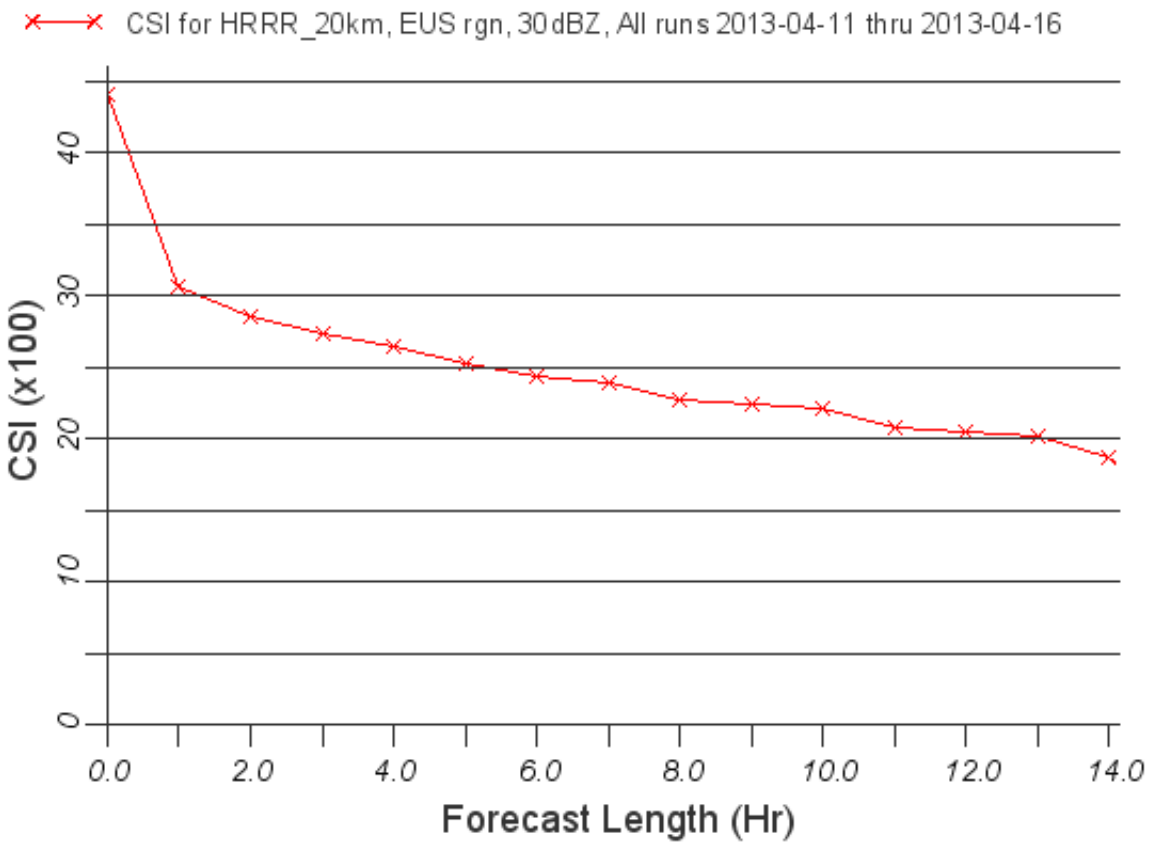


Fig. 3 30 dBZ reflectivity Critical Success Index (CSI) skill score up-scaled to a 20-km grid as a function of forecast length for the frozen 2013 GSD RAP / HRRR configuration currently running in real-time Verification for the period from 11 – 15 April 2013.

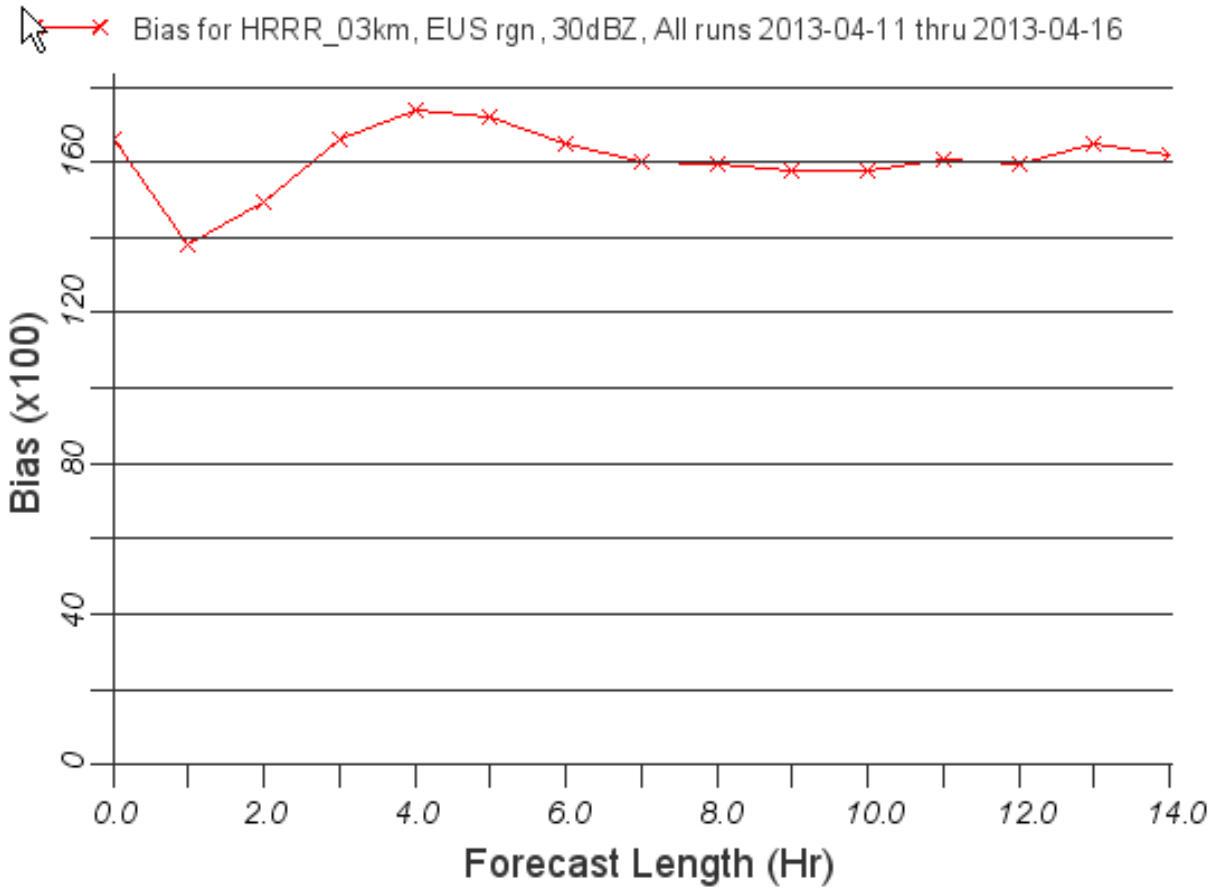
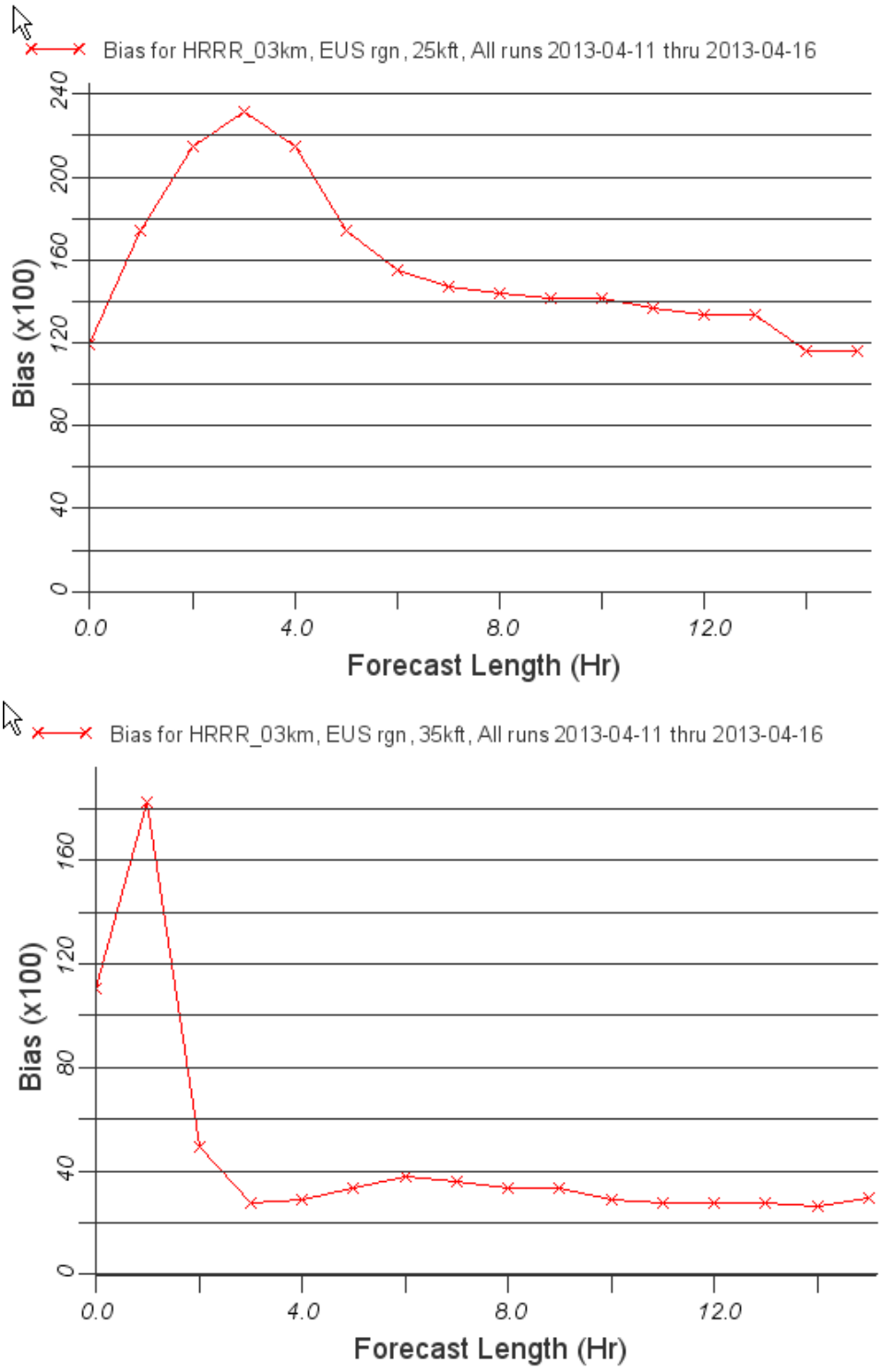


Fig. 4 30 dBZ reflectivity bias score as a function of forecast length for the frozen 2013 GSD RAP / HRRR configuration currently running in real time. Verification for the period from 11 – 15 April 2013.

Finally, in Figs. 5 and 6,we show two very preliminary plots of echo-top bias as a function of forecast length on the native 3-km domain. First, in Fig. 5, we see that for the lower echo-top height of 25 kft., after an initial period of high bias for the first few hours, the bias settles to a value slightly in excess of 1. In contrast, fig. 6 shows that for the higher threshold of 35 kft., after a similar period of high bias in the first few hours, the bias settles to a low value of ~ .35. It is important to note that the sample size is quite small for the very sensitive forecast parameter. In a later update to this report, we will present more robust statistics and comparisons with the 2012 configuration for both echo-top and reflectivity.



Figs. 5 and 6 Echo-top bias score on the native 3-km grid for (top) 25 kft threshold and (bottom) 35kft threshold as a function of forecast length for the frozen 2013 GSD RAP / HRRR configuration currently running in real time. Verification for the period from 11 – 15 April 2013.